A Vision System for an Unmanned, Non-lethal Weapon

Greg Kogut, Larry Drymon

Space and Naval Warfare System Center, San Diego, Code 2371, 53406 Woodward Road,
San Diego CA 92152-7383, USA;

ABSTRACT

Unmanned weapons remove humans from deadly situations. However some systems, such as unmanned guns, are difficult to control remotely. It is difficult for a soldier to perform the complex tasks of identifying and aiming at specific points on targets from a remote location. This paper describes a computer vision and control system for providing autonomous control of unmanned guns developed at Space and Naval Warfare Systems Center, San Diego (SSC San Diego).

The test platform, consisting of a non-lethal gun mounted on a pan-tilt mechanism, can be used as an unattended device or mounted on a robot for mobility. The system operates with a degree of autonomy determined by a remote user that ranges from teleoperated to fully autonomous.

The teleoperated mode consists of remote joystick control over all aspects of the weapon, including aiming, arming, and firing. Visual feedback is provided by near-real-time video feeds from bore-site and wide-angle cameras. The semi-autonomous mode provides the user with tracking information overlayed over the real-time video. This provides the user with information on all detected targets being tracked by the vision system. The user uses a mouse to select a target, and the gun automatically aims the gun at the target. Arming and firing is still performed by teleoperation. In fully autonomous mode, all aspects of gun control are performed by the vision system.

Keywords: computer vision, unmanned, non-lethal, weapon

1. INTRODUCTION

Commanders charged with the decision of whether or not to use non-lethal measures have to strike a balance between three objectives: force protection, mission accomplishment, and the safety of non-combatants.[1] Force protection can be difficult to ensure while using non-lethal weapons, particularly in environments with rapidly varying threat levels and unknown or unidentified combatants. Robotic delivery of NLWs effectively solves the force protection problem by moving personnel to a safe standoff distance.

Non-lethal weapons (NLWs) have the potential to play a large future role in the military and police applications. However, according to the National Research Council (NRC), NLWs have yet to be fully adopted because of several technological shortcomings.[2] Two of the recommendations of the NRC to overcome these shortcomings are to "accelerate technology programs that explore the creative use of remotely piloted and robotic vehicles to deliver NLWs," and "expand efforts to develop, improve, and better utilize existing sensor technologies for non-lethal weapons applications."[2]

SSC San Diego had developed a networked, remotely operated paintball gunpod which addresses both of these issues, and serves as a test bed for exploring robotic and sensor development for NLW delivery. The gunpod is digitally networked, and can act either as a standalone or robot-mounted weapon. The gunpod is designed to offer three modes of operation: teleoperated, semi-autonomous, and autonomous. This paper discusses the hardware, software, and control algorithms used in the SSC San Diego gunpod.

Report Documentation Page

1. REPORT DATE **2004**	2. REPORT TYPE	3. DATES COVERED **00-00-2004 to 00-00-2004**

4. TITLE AND SUBTITLE **A Vision System for an Unmanned, Non-lethal Weapon**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Space and Naval Warfare System Center,Code 2371,53406 Woodward Road,San Diego,CA,92152-7383**	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
SPIE Proc. 5608: Intelligent Robots and Computer Vision XXII, Philadelphia, PA, 25-27 October 2004

14. ABSTRACT
Unmanned weapons remove humans from deadly situations. However some systems, such as unmanned guns, are difficult to control remotely. It is difficult for a soldier to perform the complex tasks of identifying and aiming at specific points on targets from a remote location. This paper describes a computer vision and control system for providing autonomous control of unmanned guns developed at Space and Naval Warfare Systems Center, San Diego (SSC San Diego). The test platform, consisting of a non-lethal gun mounted on a pan-tilt mechanism, can be used as an unattended device or mounted on a robot for mobility. The system operates with a degree of autonomy determined by a remote user that ranges from teleoperated to fully autonomous. The teleoperated mode consists of remote joystick control over all aspects of the weapon, including aiming, arming, and firing. Visual feedback is provided by near-real-time video feeds from bore-site and wide-angle cameras. The semi-autonomous mode provides the user with tracking information overlayed over the real-time video. This provides the user with information on all detected targets being tracked by the vision system. The user uses a mouse to select a target, and the gun automatically aims the gun at the target. Arming and firing is still performed by teleoperation. In fully autonomous mode, all aspects of gun control are performed by the vision system.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**8**	

2. HARDWARE

2.1. Paintball gun and pan-tilt mechanism

An WDM 2001 *Angel* paintball gun is used. The *Angel* fires 10-13 round per second, and uses 0.68-caliber RPS *Marballizer* ammunition. The gun is mounted on a custom, SSC developed pan-tilt mechanism that employs a 24-volt DC *Ultra Motion Smart Actuator* for tilt actuation and a *Silvermax* motor for pan control. A protective shroud encases the gun and pan-tilt mechanism. The gunpod is shown in Figure 1.

2.2. Processing module

An embedded computer system is colocated with the gun, in a protective box. The embedded computer system digitizes video from multiple video feeds, and performs all gun control, computer vision and networking functions. The embedded computer currently consists of a PC104+ form-factor Pentium III, and a digital frame-grabber. In addition to the processor, the NLW's processing module contains a miniature Ethernet switch, an 802.11g wireless radio, and, optionally, a hardware video codec. A symbolic diagram of the contents of the computer system is shown in Figure 2. An optional battery powerpack allows the system to operate wirelessly for both data and power.

The 802.11g radio provides the NLW with a peak 54 Mbps data raate, more than sufficient for streaming multiple live video feeds, as well as control data to multiple users.

The optional hardware codec digitizes, compresses, and serves analog video other Ethernet. The *IndigoVision VideoBridge* codec compresses in either MJPEG or H.261 format. In addition, the PC104 processor is capable of performing the same function. However, the variable processing load on the main CPU can result in a variable framerate or latency, while the hardware codec provides a constant framerate, which can be important while making targetting and firing decisions from a remote location.

While currently only vision sensors are employed, the processing module has the capacity to accept other sensor modalities.

Fig. 1 SSC paintball gunpod with protective shroud.
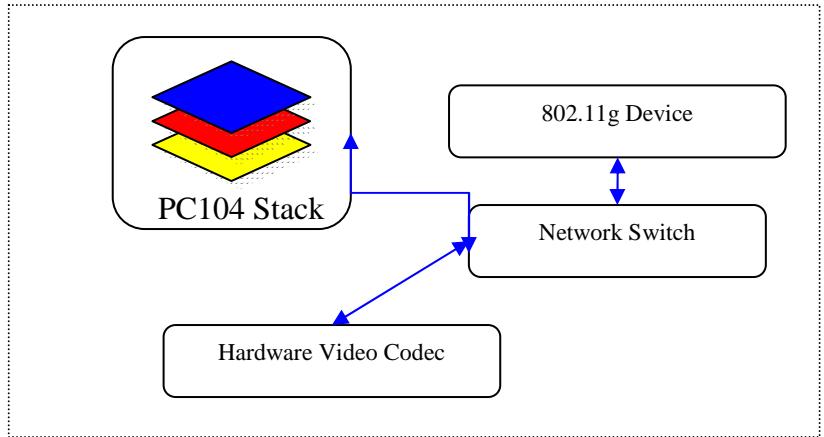
Fig. 2 Data flow within the NLW processing module.

2.3. Sensors

Initial development and testing was performed using an omnidirectional visual sensor for low cost and ease of development. More sophisticated sensors, such as scanning laser or infrared may be easily added to the current architecture. The omnidirectional camera consists of a hyperbolic mirror which collects light over 360 degrees and focuses it onto a conventional CCD. The center axis of the 360-degree field-of-view is placed at or near the axis of pan rotation on the pan-tilt platform. This close placement minimizes error due to parallax. This setup is very inexpensive, and requires no calibration if the assumption is made that all tracked targets are touching a planar surface (flat ground). The prototype sensor installation and its relationship to the gun pod are shown in Figure 3. Conventional cameras, or cameras not co-located with the gun platform, may also be used, but require a rigorous camera calibration upon setup, such as Tsai's technique for camera calibration.[3] Other types of sensors also have the potential to improve the range and accuracy of the system.
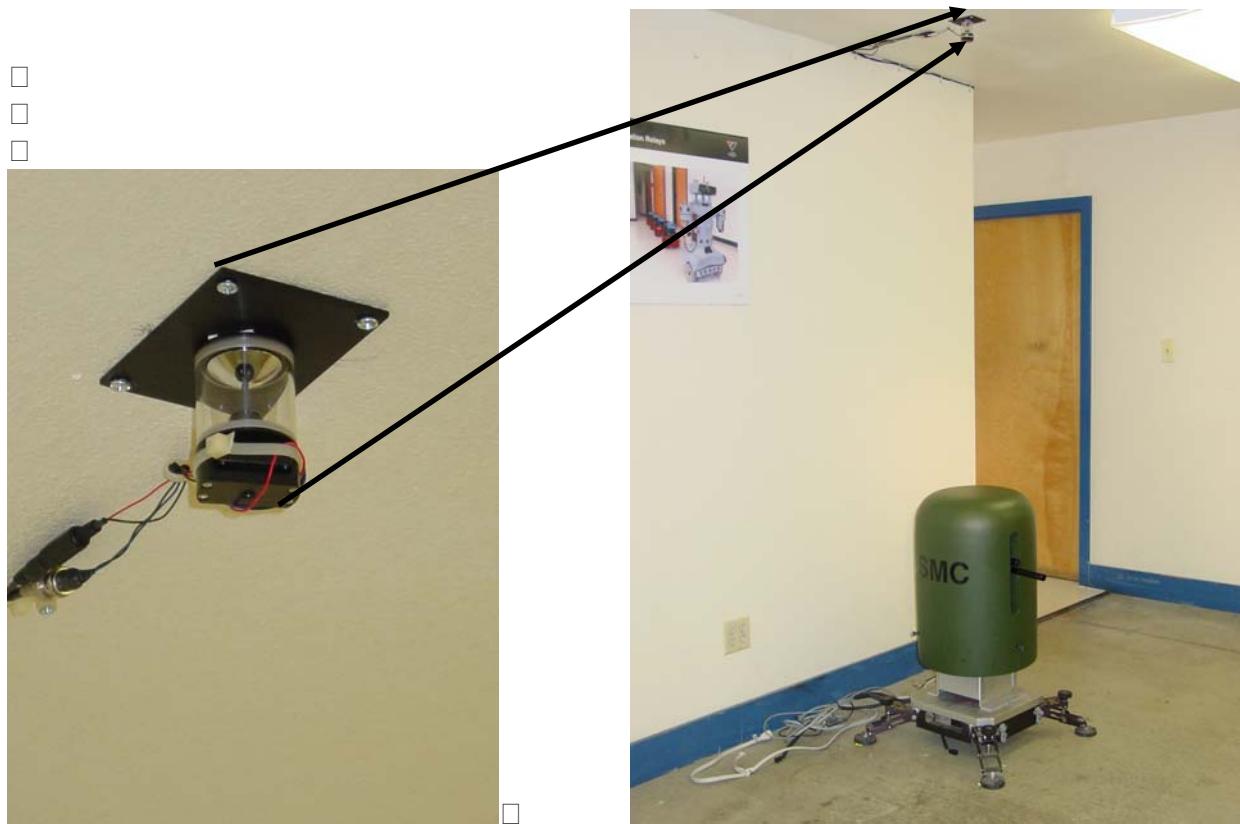
Fig. 3 The omnidirectional sensor (left), and its mounting position above the pan axis of the gun pod.

The omnidirectional sensor, mounted as shown, has an approximate effective range of 25m and an effective area of approximately 1900m^2 However, initial testing was performed under shorter ranges, and did not test the limits of the sensor's range or the range of the NLW.

□

3. SOFTWARE ARCHITECTURE

This section describes a solution for short range visual tracking of multiple simultaneous moving objects. In this application, a visual sensor is used.   The output of the system are control parameters sufficiently accurate to quickly and effectively cue a motorized pan-tilt platform to any of the tracked objects. Methods of target prosecution are also explored, an important topic for efficient use of a NLW when encountering complex situations, such as a large crowd.

The sequence of operation is: data acquisition, motion detection, multi-target tracking, and target selection. Each of these steps is discussed in detail below.

3.1. Data Acquisition

Data arrives from the sensor as a standard NTSC analog video signal. The data is digitized by a PC104 form-factor video digitizer board, then downsampled to a 320x240 array of RGB pixels. The RGB color space is used in this initial implementation, however other color spaces may produce better results in some types of computer vision.   An example image is shown in the first image of Figure 4.

3.2. Motion Detection

The motion detection scheme used is similar to, and derived from, those described by Hong and Hongbin[4] and Duckett[5]. The motion detection algorithm both detects movements, and calculates several features about all detected motion which is subsequently used during the tracking phase. The second image in Figure 4 shows an example of the results of motion detection.

3.2.1.   Background Estimation

Motion is detected by the background subtraction method. However, changes in the "background" such as lighting changes or moved furniture should not be classified as detected motion over the long term. Therefore, a statistical background model is used. Each color channel (R,G,B) of each image pixel is stored as a mean and variance (μ, σ^2) over a predefined time period.   The background model is updated recursively by each incoming video frame. Details are discussed in Hong and Hongbin.[4] This allows the background to "absorb" changes and varying lighting conditions.

3.2.2. Motion Region Detection

If any of the three color channel, R, G, or B varies by three or more standard deviations from that defined by the background model, the corresponding pixel is defined as part of a detected motion.   The result is of this stage is a binary image which contains regions of detected motion. This detection is often subject to high levels of noise, so morphological filters (erosion and dilation) are applied to "clean up" the binary image. A connected components algorithm is also run so that all connected pixels are classified as part of the same object.

3.2.3. Feature Calculation

A number of object features are calculated about each moving object at this stage. These include

- Size – number of pixels occupying the object
- Color – the mean and variance of the R, G, B values comprising the object

- Shape – the detected region is used to feed an ellipse-fitting algorithm, and the resulting major and minor axes of the ellipse are recorded as an approximation of object shape

Fig. 4 An example sequence showing three people being tracked through a cluttered lab area. From left to right: raw input image; motion detection image showing three detected people; tracking and targetting vector overlaid over each of the people in the image

3.3. Tracking

Motion information is fed to a Kalman Filter based tracking system. The Kalman Filter is a mainstay of tracking system, and effectively minimizes errors or noise produced in the image capture and motion detection steps of the system. The Kalman Filter is very fast, and easy to implement in embedded hardware. The Kalman Filter allows for predictive calculation, allowing a gun platform to be aimed at the expected future position of a tracked target. This prediction is generally not needed in practice, however, since the update rate of the vision system is generally more than fast enough to keep up with most pedestrian or vehicular motion without the need of predictive feedback.

The implemented Kalman filter is similar to that used in Duckett.[5] The filter is used to target position and velocity, and the velocity is used to predict the location of the target in the next time step. Separately, the features calculated during the detection phase are also tracked as constants, giving a statistically optimum estimation of the target features at any time step.

Solving the data association problem is a key to successful implementation of a Kalman filter tracking system. The data association problem is the association of newly detected "blobs" with already established tracks. Incorrect data association can lead to false alarms and erroneous track data. A statistical method is used to determine the likelihood that each "blob" should be associated with a given track. First, the blob must fall within the envelope area predicted by the Kalman filter. This represents the area the tracked target could possibly have traveled during a time step, given its position and velocity as determined by the filter model. Second, the features calculated during the motion detection phase are formed into a feature vector. The Mahalanobis distance between this vector, and a vector formed from the feature set tracked by the Kalman Filter is calculated. This distance ensures that the size, color, and shape of the blob are similar to the same predicted features for an object being tracked. If a blob is within the predicted envelope, and the calculated Mahalanobis distance is not greater than a predetermined threshold, than the blob is assigned as continuation of an establish track.

The output of the tracking phase is a list of unique moving objects within range of the sensor, as well as their relative locations and velocities within the image-space of the camera. This output is sufficient to drive the NLW gun platform to aim at any given target.

3.4. Target Selection

A selected target is one which the pan-tilt mechanism is actively tracking. Target selection in the NLW's fully autonomous mode becomes a problem when there is more than one target being actively tracked. Schemes for choosing among targets are largely application dependent. The current architecture has been implemented with several different target selection algorithms, including: a) select closest target first, b) select target closest to a predefined position, c) select quickest moving target, and d) select target which requires least movement along the pan-tilt axes of the NLW.

4. TESTING

The NLW is still in an initial development phase, and has yet to undergo rigorous testing. Testing the tracking system consists of running the targeting system offline using a suite of test video for which manually calculated ground truth tracking data exists. The output of the tracking system is compared to the ground truth. Testing, at the time of writing, has only occurred at short ranges of <20m. The NLW system shows sufficient accuracy at these ranges for accurate firing. Live firing at moving targets has not been undertaken yet.

5. USER INTERFACE

User interface design is very important for remote control of an NLW. It is important for any user to have a clear, reliable awareness of both where the gun is aiming, as well as awareness of the area within the sensor range of the gun. This information is provided to the user through the use of multiple, independent sensors.

A bore-sight camera physically mounted on the gun barrel provides a view of where the gun is currently aiming. This video stream can pass through a hardware codec which is independent from the rest of the computational hardware. Latency from this hardware codec depends largely on the medium of wireless transmission. Transmission using conventional 802.11b radios is typically on the order of a few tenths of a second, but cannot be guaranteed due to the nature of the TCP/IP protocol. True real-time transmission, however, can be achieved through the use of a protocol with guaranteed quality of service, or through a wired link.
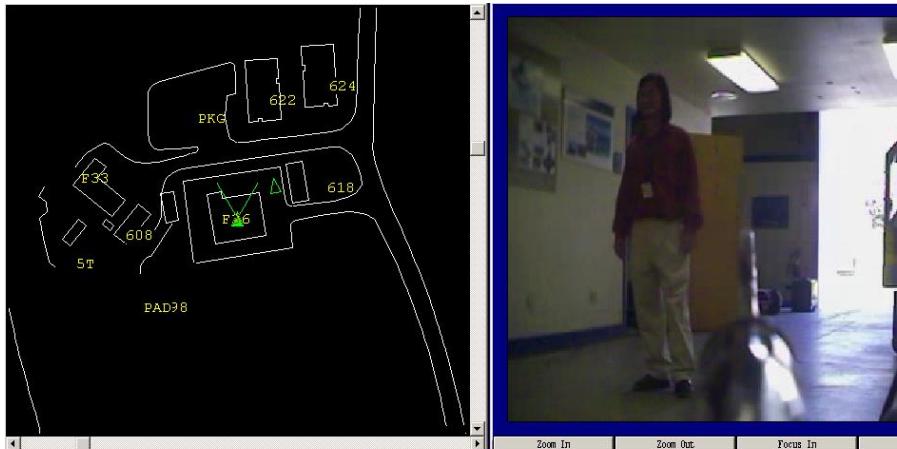
Fig. 5 Teleoperation user interface. The image on the right shows live bore site video. The gun barrel can be seen in the right half of the image. The image on the left shows an orthographic map of the area with the sensor range of the gun. The camera view range of the bore site camera is represented by a cone in the center of the image.

The rest of the sensors are digitized and broadcast through the embedded computer. The embedded software can serve the data in multiple formats, and can overlay various types of information, such as velocity, target-type, currently targeted object, etc. Sensor data from the embedded computer is generally has more latency than the bore-sight camera because software codecs are used. However, the latency is still generally on the order of a few tenths of a second using standard 802.11b radios. Similar to the bore sight camera, better performance can be achieved if required by the application.

Figure 5 shows the manual, teleoperated mode of the NLW. A joystick control gun motion, arming, and firing. Feedback is provided by near-live bore site video. An orthographic map of the area surrounding the gun with icons showing the NLW as well as targets provides more information to the user.

A fully autonomous mode, which doesn't need an interface, is also available. This mode automatically aims at targets according to one of the algorithms described above. For safety purposes, arming and firing is still performed via teleoperation in this mode, but need not be depending upon the application.

The next stage in user interface design will be to add a semi-autonomous interface. This interface will allow user-directed target designation by clicking on a target in live video, or on an iconic representation of a target in a map. The NLW will then automatically track the given target. This interface takes the burden of accurate aiming off human users, and also overcomes control difficulties which arise from the latencies common in digital communications. However it still leaves the decision to arm and fire entirely with human users, which is important in many applications.

6. MOBILE USE AND FUTURE WORK

In addition to use a standalone NLW, the gun pod can also be easily mounted on mobile platforms. This capability allows the NLW to be placed in potentially hostile or dangerous environments without risk to personnel. Figure 6 shows the NLW mounted on a mobile robot.

Future work includes robust testing of the sub-systems and exploration of longer-range sensor modalities, such as radar and scanning laser. The ideal sensor should extend beyond the effective range of the weapon so that the capabilities of the NLW are maximized.

Other work includes tailoring user interfaces to specific applications. Particularly, the semi-autonomous interface needs to allow personnel to confidently and easily make decisions about whether or not to fire.

Fig 6. NLW mounted on a large exterior robot.

8. MOBILE USE AND FUTURE WORK

Use of non-lethal weapons often occurs in situations which can quickly escalate to a point where deadly force may be needed. Personnel in such situations are often reluctant to use NLWs because they hamper their ability to be able to respond with deadly force, if needed. Using robotic delivery systems for the delivery of NLWs removes personnel from the situation, and solves the force-protection problem. The work describes in this paper represents a first step towards effectives use of robotic NLWs.

REFERENCES

1. Department of Justice and Department of Defense Joint Technology Program*: Second Anniversary Report, NIJ Research Report*, February 1997.
2. Committee for an Assessment of Non-Lethal Weapons Science and Technology, National Research Council, "An Assessment of Non-Lethal Weapons Science and Technology," Washington, *National Academies Press*, 2003. 198 p.
3. Tsai, Roger Y. (1986) An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision, *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Miami Beach, FL, 1986, pp. 364 374.
4. Hong Liu, Wenkai Pi, Hongbin, Zha, "Motion Detection for Multiple Moving Targets by Using an Omnidirectional Camera," *IEEE Signal Processing Magazine*, 19(2), 2002.
5. Ducket, T., "Appearance-based Tracking of Persons with an Omnidirectional Vision Sensor ," Proc. Omnivis 2003, *Fourth IEEE Workshop on Omnidirectional Vision*, Madison, Wisconsin, June 21, 2003